Behavioral/Cognitive

The Dorsal Medial Prefrontal Cortex Responds Preferentially to Social Interactions during Natural Viewing

Dylan D. Wagner,[1] William M. Kelley,[2] James V. Haxby,[2] and Todd F. Heatherton[2]

[1]Department of Psychology, The Ohio State University, Columbus, Ohio 43210, and [2]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755

Humans display a strong tendency to make spontaneous inferences concerning the thoughts and intentions of others. Although this ability relies upon the concerted effort of multiple brain regions, the dorsal medial prefrontal cortex (DMPFC) is most closely associated with the ability to reason about other people's mental states and form impressions of their character. Here, we investigated this region's putative social category preference using fMRI as 34 participants engaged in uninstructed viewing of a complex naturalistic stimulus. Using a data-driven "reverse correlation" approach, we characterize the DMPFC's stimulus response profile from ongoing neural responses to a dynamic movie stimulus. Results of this analysis demonstrate that the DMPFC's response profile is dominated by the presence of scenes involving social interactions between characters. Subsequent content analysis of video clips created from this response profile confirmed this finding. In contrast, regions of the inferotemporal and parietal cortex were selectively tuned to faces and actions, both features that often covary with social interaction but may be difficult to disentangle using standard event-related approaches. Together, these findings suggest that the DMPFC is finely tuned for processing social interaction above other categories and that this preference is maintained during unrestricted viewing of complex natural stimuli such as movies.

Key words: medial prefrontal cortex; mentalizing; narrative; neuroimaging; social cognition; theory of mind

---

Significance Statement

Recently, studies have brought into question whether the dorsal medial prefrontal cortex (DMPFC), a region long associated with social cognition, is specialized for the processing of social information. We examine the response profile of this region during natural viewing of a reasonably naturalistic stimulus (i.e., a Hollywood movie) using a data-driven reverse correlation technique. Our findings demonstrate that, during natural viewing, the DMPFC is strongly tuned to the social features of the stimulus above other categories. Moreover, this response differs from other areas with previously well characterized response profiles such as the lateral and medial fusiform gyrus. These findings suggest that this region's dominant function in everyday situations is to support reasoning about the thoughts and intentions of conspecifics.

Introduction

Observing people interacting, be it in everyday life or when viewing movies, involves an array of perceptual and cognitive operations from recognizing faces and understanding the intent of actions to inferences concerning the thoughts and intentions lurking in other people's minds. Neuroscientific models of social

cognition commonly implicate the dorsal medial prefrontal cortex (DMPFC) in reasoning about other people's mental states (Mitchell, 2008; Spreng et al., 2009; Van Overwalle, 2009; Wagner et al., 2012). Moreover, this region is associated not only with accuracy during social cognition (Zaki et al., 2009; Moran et al., 2012; Spunt and Adolphs, 2015), but has also been shown to play a causal role in forming impressions of people (Ferrari et al., 2016).

Although a small number of studies have attempted to examine the DMPFC's role in reasonably naturalistic tasks such as viewing images of familiar faces or of social interactions (Gobbini et al., 2004; Iacoboni et al., 2004; Spiers and Maguire, 2006; Wheatley et al., 2007; Gobbini et al., 2011; Wagner et al., 2011b), the majority of prior research rests on controlled experiments in which aggregate cortical responses to a restricted number of stimulus categories are measured under a limited set of task conditions. One alternative approach, more common in vision neu-

roscience, is to instead map the response profile of a cortical region to a large number of stimulus categories in order to make inferences as to the underlying cognitive processes and stimulus dimensions encoded in brain activity (Wu et al., 2006; Vul et al., 2012; Mur et al., 2012). For example, studies show that the inferotemporal cortex supports a broad encoding of features related to animacy (Haxby et al., 2001; Connolly et al., 2012), with topographical peaks in the lateral fusiform corresponding to a category preference for highly animate stimuli such as faces (Haxby et al., 1996; Kanwisher et al., 1997; Ishai et al., 1999).

In the present study, we adopt an approach used previously by Hasson et al. (2004) which implemented a novel reverse correlation analysis that inverts the typical fMRI analysis by instead using neural activity measures in three cortical regions (i.e., lateral fusiform, medial fusiform, and intraparietal sulcus) to identify segments of the stimulus corresponding to statistical peaks in the time series of brain activity in each area. The results of this study showed that, during natural uninstructed viewing, these regions demonstrated a category preference for specific features of the stimulus. For instance, the lateral fusiform displayed a category preference for faces, the medial fusiform for outdoor locations, and the intraparietal sulcus for manual actions and tool use. Together, these findings provided strong converging evidence for experiments on more restricted stimulus spaces and with more artificial tasks.

In the present study, we characterized the response profile of the DMPFC during uninstructed viewing of the first 30 min of a popular Hollywood movie. Using a reverse correlation analysis, we extracted and characterized scenes in the movie that evoked the strongest and weakest responses in the time course of the DMPFC. To ensure the reliability of our method, we also sought to replicate the findings of Hasson et al. (2004) in the lateral, medial fusiform, and intraparietal sulcus. We hypothesized that, even in the absence of task demands, the DMPFC's response profile would nevertheless be dominated by movie scenes depicting social interactions between characters. Although the present data-driven approach necessarily requires surrendering a degree of experimental control, it allows for inferences beyond experiments with limited sets of categories and offers a more ecologically valid test of this region's role in social cognition under reasonably naturalistic conditions (for a similar argument, see Zaki and Ochsner, 2009; Hasson and Honey, 2012).

Materials and Methods

Participants. The imaging portion of the study included 34 right-handed participants (27 female; mean age 22.2 years) with normal or corrected-to-normal visual acuity and no abnormal neurological history (Wagner et al., 2011a). The ratings of response profile videos portion of the study included 132 participants (90 female; mean age 19.3 years) recruited from introductory psychology courses at Dartmouth College (see "Content analysis" section below). Participants gave informed consent in accordance with the guidelines set by the Committee for the Protection of Human Subjects at Dartmouth College.

Stimuli. Stimuli consisted of the first 30 min of footage from the movie *Matchstick Men* (2003). This movie had the desirable characteristic of being composed primarily of scenes containing the protagonist (played by Nicholas Cage) and one other character (i.e., dyads), the protagonist alone, or no characters at all (e.g., shots of scenery and location establishing shots). Because of scanner run-length limitations, the film was edited into three 10 min segments each preceded and followed by 60 s of null event trials consisting of a white fixation cross against a black background.

Movie segmentation. The stimulus movie was segmented into different scene types using the video annotation tool Anvil (http://www.anvil-software.org/). Scenes were segmented according to the following simple

rule: social scenes were defined as those in which the protagonist and one other character were present (visible or implied through conversation), single character scenes were those in which the protagonist was alone on screen, and people absent scenes were those in which no characters were visible. In this way, every frame of the film was assigned to one of these three conditions.

Procedure. Participants were informed that they were taking part in a study on the neural correlates of movie watching and were instructed to pay attention to the film. Participants listened to the audio soundtrack of the film via a pair of MRI-compatible headphones. After scanning, participants were debriefed and completed a measure of narrative transportation (Dal Cin et al., 2004), the interpersonal reactivity index (Davis, 1983), and a set of short questions assessing whether participants paid attention to the film.

Image acquisition. MRI was conducted with a Philips Achieva 3.0 tesla scanner using an eight-channel phased array coil. Functional images were acquired using a T2*-weighted echoplanar sequence (TR: 2500 ms; TE: 35 ms; flip angle: 90°; field of view: 24 cm). For each participant, 3 runs of 288 whole-brain volumes (30 axial slices per whole-brain volume, 4.5 mm thickness, 0.5 mm gap, 3 × 3 mm in-plane resolution) were collected.

Image preprocessing. Preprocessing was conducted using SPM8 (Wellcome Department of Cognitive Neurology, London) in conjunction with a suite of in-house tools for preprocessing and analysis (available at: http://github.com/wagner-lab). First, images were corrected for differences in acquisition time between slices and then realigned within and across runs via a rigid body transformation in order to correct for head movement. Images were unwarped to reduce residual movement-related image distortions not corrected for by realignment and then normalized into a standard stereotaxic space (3 mm isotropic voxels) based on the SPM8 EPI template that conforms to the ICBM 152 brain template space (Montreal Neurological Institute). Normalized images were spatially smoothed (6 mm full-width-at-half-maximum) using a Gaussian kernel to increase the signal-to-noise ratio (SNR). Volumes were inspected for scanner- and motion-related artifact based on an examination of the realignment parameters and temporal SNR for each run in each participant.

ROI selection. Analyses were conducted using spherical ROIs (6 mm radius) centered on coordinates for the DMPFC and for three regions that have a well characterized stimulus category preference, namely: the right lateral fusiform area for faces, the medial fusiform area for places and scenes, and the left intraparietal sulcus for actions. Except for the DMPFC, the remaining regions are approximately the same as those used in prior work by Hasson et al. (2004) and were chosen both to ensure that our method was consistent with that study's findings and to serve as comparison regions for the DMPFC (e.g., to demonstrate discriminative validity). Coordinates for these a priori ROIs were selected from a large-scale automated meta-analytic database of neuroimaging studies (Yarkoni et al., 2011) identifying peak coordinates in the reverse inference maps for the terms mentalizing (98 studies), fusiform face area (61 studies), place (106 studies), and actions (397 studies). These coordinates were rounded to the nearest voxel in our 3 × 3 × 3 mm space.

Reverse correlation analysis. A data-driven reverse correlation was calculated between the stimulus movie and the BOLD time series in each of the four regions described above. For each ROI, a mean time series was extracted and adjusted for covariates of no interest (motion parameters, linear trends to make the data stationary, and session means to allow for concatenation of different sessions). To minimize the impact of any signal shared across brain regions, we also adjusted each region's time series by regressing out the mean time course across all visually active voxels based on a contrast of movie versus fixation baseline ($p < 0.05$ corrected). For each subject and each ROI, we removed time points corresponding to the resting fixation that began and ended each functional run and an additional 14 volumes to cover any transient bold responses that might arise from switching between rest and film. Within each ROI, the time series was standardized, temporally smoothed with a moving average of 3 time points, and shifted backwards in time by 2 TRs (i.e., 5 s) to account for the time to peak of the hemodynamic response function (Miezin et al., 2000). Next, a *t* test across subjects was performed on every time point in each ROI, resulting in a vector of *t* statistics corresponding

to reliable peaks and troughs in the time series. To correct for multiple comparisons across separate time points, a false discovery rate (FDR) correction was calculated for each ROI (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001) using a q value of 0.05. Finally, a peak identification algorithm was used to identify local peaks and troughs in the time series (minimum peak to peak width of 3 TRs) and the movie frames corresponding to these were extracted. To aid in visualization and content analysis of these scenes, a second movie clip extraction was conducted for the 10 highest-ranking peaks and troughs.

To determine the reliability of the reverse correlation procedure described above, we calculated the reverse correlation for 1000 simulations of splitting the participants into one of two groups with the constraint that each grouping of participants is unique. For each group, we calculated the average correlation between t-statistic vectors, as well as the percentage overlap between peaks detected in each group, by the reverse correlation procedure, at both uncorrected and corrected thresholds. Finally, we computed the distribution of scene types based on the movie segmentation for each group across the 1000 iterations.

In addition to the ROI-based reverse correlation analysis outlined above, we also implemented a whole-brain voxelwise reverse correlation analysis following the same steps. For each voxel, we calculated the adjusted percentage of each scene category in the reverse correlation profile and then tested each response profile against the total distribution of scene types in the movie using the χ^2 goodness of fit test. To visualize the response profiles, we calculated a difference score among social, single, and person absent scenes and masked the resulting statistical parametric map by the whole-brain-corrected map of χ^2 statistics ($\chi^2 > 28.94$, $p < 0.05$, Bonferroni corrected). For visualization, the resulting maps were overlaid onto inflated cortical renderings using Caret (Van Essen et al., 2001) and a minimum cluster surface area of 50 mm.

Intersubject correlation reliability analysis. Intersubject correlations (ISCs) for each ROI were calculated by taking the mean correlation $r(k, j)$ between the time series for each participant (k) and the mean time series across all $n − 1$ participants (j) (Lerner et al., 2011). To assess the reliability of the ISC for each region, we generated Monte Carlo simulations of the empirical distribution of correlations under the null hypothesis of no ISC. To generate null data, the time series underwent one of two possible forms of temporal randomization. In the first method, the time series of the n-1 group average is circularly shifted by random amounts such that the two time series no longer refer to the same segments of the stimulus and the expected value of $r(k, j) = 0$. A second method "scrambles" the phase of the time series of the n-1 group average (j) in the frequency domain. This is accomplished by performing a Fourier transform of the original time series and applying the phase structure of random noise while maintaining the spectral amplitude of the original time series. This phase-randomized time series is transformed back to the time domain and scaled to the range of the original time series. Both methods allow for the generation of synthetic null data that maintain most characteristics of the original time series (e.g., spectral amplitude, temporal auto-correlation, and range) except for the phase and in practice both give nearly identical distributions. Here we report the results based on the phase-scrambling procedure.

To construct a distribution of null-effect statistics, the average ISC for each ROI was calculated using 10,000 iterations of phase-scrambled null data. Examination of these empirically derived distributions revealed that they were well approximated by a normal distribution. Therefore, for each ROI, a normal distribution was fitted to the empirically derived distribution and used to derive the mean and SD of that distribution. Using these parameters, a z test was performed to test the significance of the mean ISC for each ROI.

Content analysis of reverse correlation response profiles. Two approaches were used to characterize the content of each region's response profile. The first involved labeling significant peaks and troughs with the category membership for that time point as determined by movie segmentation (e.g., social, single character, or characters absent). From these labeled time points, the percentage of significant peaks and troughs corresponding to each scene category was calculated and adjusted for the overall frequency of each category across the movie.

The second approach attempted to rule out any putative confirmation bias by asking a group of "neuroscientifically naive" participants ($N = 132$) to characterize the response profile videos. Participants viewed four 25 s videos corresponding to the top 10 peak volumes for each ROI. Participants were informed that the video clips were created based on brain activity from people watching a movie and that these clips represented scenes that different brain regions "preferred" the most. After viewing all of the clips, participants viewed them a second time, this time writing down three verbs to describe the content of each clip. Finally, after a third viewing, participants were asked to rate the response profile videos according to the degree to which the videos exemplify the following dimensions: social interactions, faces and facial expressions, actions, objects, and scenes/places. Ratings were conducted on a five-point Likert scale (1 = not at all, 3 = somewhat, 5 = extremely). This order of presentation prevented the rating scale dimensions (e.g., faces, actions etc.) from biasing participants' generation of verbs to describe the reverse correlation video clips. Therefore, this allowed us to characterize the content of the video clips based on naive participants' use of verbs as well as from explicit ratings along dimensions thought to underlie the category preference of these different brain regions. Verb descriptions for each response profile were preprocessed to account for differences in verb tense (e.g., eat and eating are considered the same verb) and verb frequencies were converted to a measure of distinctiveness (e.g., how often a verb was reported for given video clip after controlling for its overall frequency across all response profile videos).

Results

Reverse correlation analysis

A data-driven reverse correlation analysis was performed on BOLD time series in the DMPFC and in three regions with well characterized stimulus category preferences (i.e., the lateral and medial fusiform and intraparietal sulcus) to examine features of movie scenes that contributed to the response profile of each region. To facilitate visualization and interpretation, the region specific profiles described below and in Figure 1 focus on only the top 10 highest-ranking peaks and troughs for the DMPFC (Fig. 1A) and three additional regions (Fig. 1B–D). Characterization of the complete set of significant peaks and troughs for each region ($p < 0.05$, FDR corrected) is available in Table 1 and is further described in the next section.

These results replicated known category preferences with the lateral fusiform demonstrating a preference for close-up shots of faces (Fig. 1B), the medial fusiform preferring outdoor scenes and scenes containing multiple objects (Fig. 1C) and the left intraparietal sulcus preferring scenes involving actions, particularly manual actions such as eating, smoking a cigarette, or washing hands (Fig. 1D). In contrast, the DMPFC response profile was dominated by scenes involving interactions between characters (Fig. 1A). In addition, this category preference was specific to peak responses; response profiles derived from troughs demonstrated a near complete absence of social interactions and were composed primarily of people alone or of no persons at all.

Region-specific response profiles for the other three regions also showed discriminant validity within regions, with response profiles for significant troughs in the time course demonstrating a near absence of each region's putative preferred category. For instance, the troughs in the lateral fusiform contain very few close-up shots of faces (Fig. 1B, bottom row), whereas the troughs for the medial fusiform predominantly involve close-up scenes of people (Fig. 1C, bottom row).

Proportion of scene types in reverse correlation video clips

To better specify the content of region-specific response profiles, we calculated the percentage of scene types (social interactions,
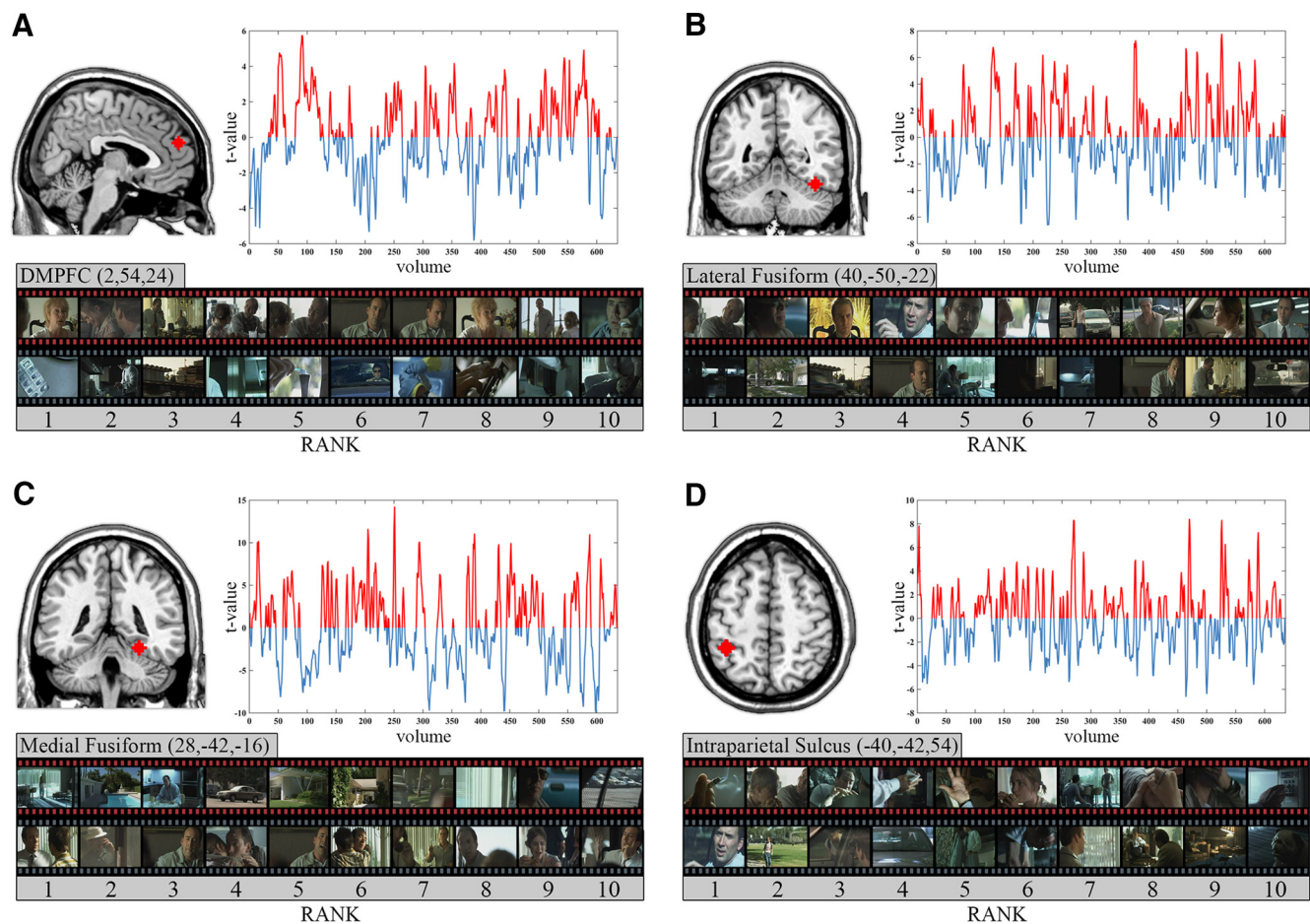
Figure 1. Reverse correlation analysis in the DMPFC (**A**), the lateral fusiform (**B**), the medial fusiform (**C**), and the intraparietal sulcus (**D**). The top right plot for each region demonstrates the volume-wise *t*-statistic for a one-sample *t* test across subjects used to calculate significant positive (peaks in red) and negative (troughs in blue) responses. Motion picture frames extracted from the 10 highest peaks and troughs ($p < 0.05$, FDR corrected) for each region are presented in rank order. All coordinates are in Montreal Neurological Institute stereotaxic space.

Table 1. Reliability of the intersubject correlations for each region and the percentage of each scene category type present in the complete response profile vectors of peaks and troughs for each region

Brain region	ISC	Social	Single	Person absent	Volumes
Peaks					
Dorsal MPFC (2, 54, 24)	0.2	100%	0%	0%	22
Lateral fusiform (40, −50, −22)	0.3	38.8%	41.8%	19.4%	34
Medial fusiform (28, −42, −16)	0.5	12.1%	21.7%	66.2%	49
Intraparietal sulcus (−40, −42, 54)	0.3	31%	65.2%	3.8%	30
Troughs					
Dorsal MPFC (2, 54, 24)	—	5.7%	35.8%	58.5%	18
Lateral fusiform (40, −50, −22)	—	20.8%	22.5%	56.6%	37
Medial fusiform (28, −42, −16)	—	77.8%	22.2%	0%	45
Intraparietal sulcus (−40, −42, 54)	—	37.2%	22.6%	40.2%	35

Simulations comparing the mean intersubject correlation with a null distribution derived from phase-scrambled response time series demonstrated that the ISC was reliable for all four regions (all $Z > 15.7$, all $p < 0.00001$). All peak and trough response profiles deviated significantly from the distribution of scene types in the movie (all $\chi^2 > 14$, all $p < 0.001$). Pairwise χ^2 tests of independence showed that each region's response profile in peaks and troughs was independent of that of the other regions (all $\chi^2 > 6$, all $p < 0.04$).

character alone, characters absent) for each of the response profiles of significant peaks and troughs using the entire reverse correlation response profile. Results of this analysis showed a strong selectivity for social scenes in the peak reverse correlation vectors for the DMPFC (Table 1), whereas the troughs showed the opposite pattern. Findings in other regions were consistent with the category preference for these regions, with the lateral fusiform showing its strongest responses to scenes involving people (both social and single scenes) and the medial fusiform responding primarily to scenes in which no persons were present (Table 1).

χ^2 goodness of fit tests demonstrated that each region's pattern of category preference all deviated significantly from the total distribution of scene types in the movie (all $\chi^2 > 14$, all $p < 0.001$) and a series of pairwise χ^2 tests of independence indicated that each region's response profile was independent of that of other regions (all $\chi^2 > 6$, all $p < 0.04$).

Content analysis of response profile videos
In the previous content analysis, scenes were categorized by movie segmentation into social, character alone, or no-

Table 2. The four most distinctive verbs used to describe the content of response, profile videos along with the number of unique verbs for each profile video description

Brain region	E Verbs	Four most distinctive verbs			
		1	2	3	4
Dorsal MPFC	109	Talking (56)	Smiling (31)	Laughing (24)	Interacting (21)
Lateral fusiform	108	Twitching (88)	Staring (14)	Eating (41)	Watching (13)
Medial fusiform	120	Driving (63)	Walking (27)	Moving (26)	Playing (13)
Intraparietal sulcus	110	Smoking (75)	Eating (51)	Cleaning (23)	Washing (17)

Verb distinctiveness was calculated by taking the frequency for which each verb was reported for a particular region and adjusting that value for the overall frequency for which that verb was used across the entire corpus of verbs. Word frequency is shown in parentheses.
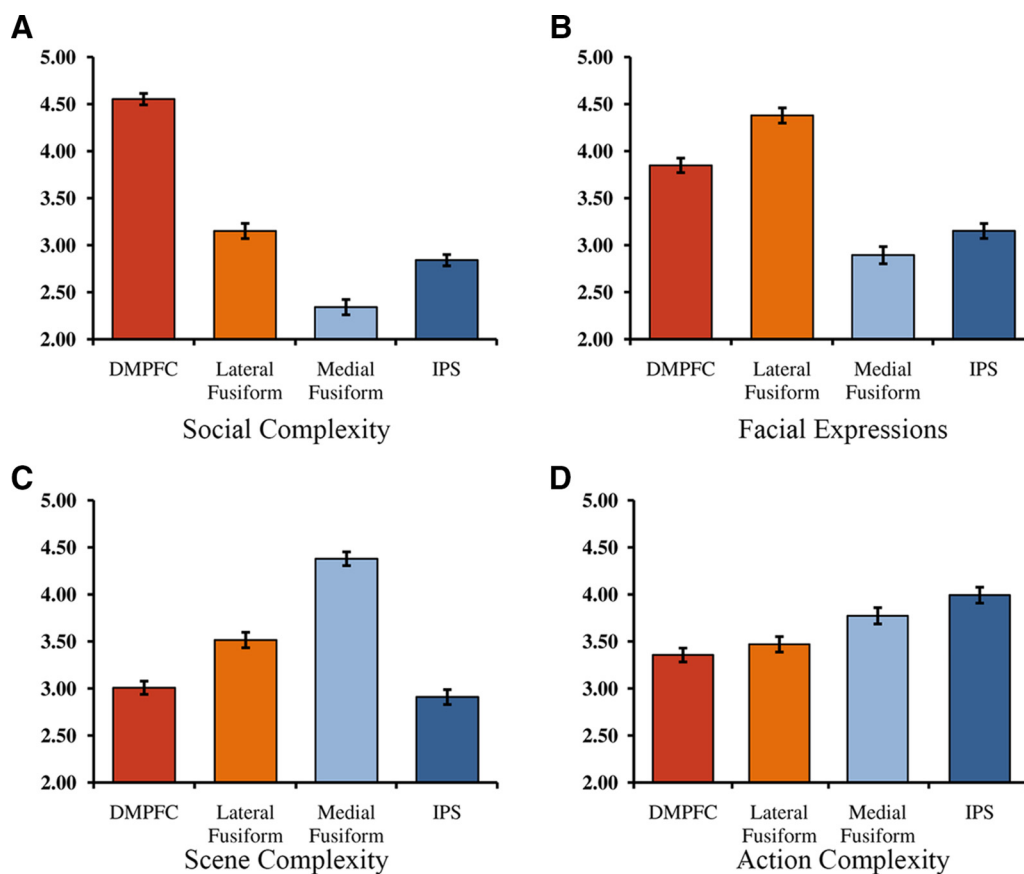

Figure 2. Ratings of response profile videos containing the 10 highest-ranking peaks along the dimensions of social complexity (*A*), facial expressions (*B*), scene complexity (*C*), and action complexity (*D*). Response profile videos were rated by an independent set of participants (*N* = 132) along dimensions related to the putative category preference of each region.

character clips. To test whether similar categorizations would be generated by individuals naive to the purpose of the study, an independent cohort of subjects (*N* = 132) were asked to generate a set of verbs and provide ratings along several dimensions for each response profile video. Examination of the four most distinctive verbs for each region revealed specific categorical differences between the response profile videos (Table 2). The verbs used to describe the content of the DMPFC video clip were all associated with social interactions (e.g., talking, smiling, interacting). In contrast, the verbs reported for the lateral fusiform, a region with a peak response profile that is similarly characterized by scenes involving people, were associated with facial movements and expressions (e.g., twitching, staring). The intraparietal sulcus was characterized by verbs having to do with manual actions (e.g., smoking, eating, cleaning). Finally, the verbs generated for the response profile video for the medial fusiform are more difficult to interpret, with most verbs relating primarily to outdoor activities (e.g., driving, walking, moving). Because participants were required to generate verbs, which are often actions performed by people, it may have been

particularly difficult to characterize the medial fusiform video which contains many scenes without any people.

During a third viewing of each video clip, subjects were asked to evaluate response profiles explicitly for the presence of social interaction, faces and facial expressions, actions, and scenes/places. Results of this analysis revealed that naive participants were sensitive to each region's putative preferred category. Specifically, participants rated video clips for the DMPFC as containing more social interactions than all other regions (all $t_{(131)} > 13.7$, $p < 0.001$; Fig. 2*A*), the lateral fusiform video clips was rated highest on facial expressions (all $t_{(131)} > 5.24$, $p < 0.001$; Fig. 2*B*), and the medial fusiform video-clip was rated highest on scene complexity (all $t_{(131)} > 8.25$, $p < 0.001$, Fig. 2*C*). Finally, although the left intraparietal sulcus was rated highest on action complexity and was significantly different from the rating for the DMPFC and lateral fusiform (both $t_{(131)} > 4.67$, $p < 0.001$), the comparison of action complexity ratings for the intraparietal sulcus

and medial fusiform was not significant ($t_{(131)} = 1.75$, $p = 0.08$; Fig. 2D).

Whole-brain reverse correlation analysis

To investigate the response profile outside of the ROIs that we selected, a whole-brain voxelwise reverse correlation analysis was performed. Because of the large number of voxels in a whole-brain analysis, we cannot visualize the response profile of each individual voxel. Nevertheless, by cross-correlating the resulting response profiles with the movie segmentation information (see Materials and Methods), we can identify voxels with response profiles that are dominated by each of the three scene categories. Results of this analysis (Fig. 3) corroborate the ROI findings by demonstrating a large cluster in the medial prefrontal cortex with a response profile that is dominated by social scenes. Similar category preferences were obtained in the temporal poles and a region of the temporoparietal junction. Two other regions, the precuneus and superior temporal sulcus, both frequently implicated in social perception, showed responses that contained a preponderance of social scenes, but to a lesser degree than the MPFC.

Reliability of ISCs and reverse correlation results in ROIs

The reliability of responses in each region was investigated in two ways. First, the reliability of the ISC between the time series in each region was computed (Table 1). Across all four regions that we investigated, the mean ISC was found to be significantly reliable compared with simulated data using phase-scrambled time series (all $Z > 15.7$, all $p < 0.00001$). Second, we examined the reliability of the reverse correlation analysis over 1000 iterations of randomly assigning participants to one of two groups and performing reverse correlation on each group separately (see Materials and Methods). The mean correlation between t-statistic vectors across these iterations demonstrated good reliability (DMPFC, mean $r = 0.59$; lateral fusiform, mean $r = 0.72$; medial fusiform, mean $r = 0.85$; intraparietal sulcus, mean $r = 0.68$). The average overlap of volumes detected by the reverse correlation procedure was generally highest in visual regions and heavily dependent on the statistical threshold used. At an α level of 0.05 (uncorrected for multiple comparisons across 636 time points), the mean overlap was 43% for DMPFC, 57% for the lateral fusiform, 79% for the medial fusiform, and 50% for the intraparietal sulcus. For FDR-corrected analyses, the mean overlap was 5.8% for the DMPFC, 53.3% for the lateral fusiform, 77.7% for the medial fusiform, and 49.2% for the intraparietal sulcus. Despite these lower values, the mean distribution of category preferences in the reverse correlation analyses was nearly identical across all 1000 iterations of splitting participants into two groups. Specifically, for the DMPFC, the mean distribution was 98.06% social, 1.93% single person, and 0% person absent for the first group and 97.95% social, 2.05% single person, and 0.007% person absent for the second group.

Discussion

Over the last 20 years, a large body of research has demonstrated a role for the DMPFC in reasoning about the mental states of others. Much of this work relies on a limited set of stimulus categories in tasks that explicitly instruct people to engage in mental state reasoning. Here, we examined the response profile of this region under reasonably naturalistic conditions (i.e., watching a movie) in the absence of any task

instructions. Results from a data-driven reverse correlation analysis based on the time series of evoked responses in the DMPFC revealed a strong category preference for social interaction scenes in the movie.

In addition, we replicated previous findings using a similar method in the lateral and medial fusiform and the intraparietal sulcus (Hasson et al., 2004). Consistent with this work, our method recovered previously described category preferences for faces, places, and hand actions in each of these regions. Content analysis of these peaks in the reverse correlation analysis for the DMPFC revealed that this region is driven primarily by scenes involving social interactions. Although the lateral fusiform also responded at least partially to social interactions scenes in the movie, its reverse correlation peaks instead showed scenes in which there was a strong percept of a face. This was corroborated by data from an independent set of naive participants asked to provide verbs and ratings to describe the response profile videos. For the lateral fusiform, these participants primarily used verbs referring to facial expressions and rated this video has having the most facial expression-related content, whereas for the DMPFC, participants described the response profile using verbs related to social interactions and rated it as having the highest complexity of social interactions of all of the reverse correlation video clips. Together, these findings replicate the prior work of Hasson et al. (2004) using a different stimulus movie and extend these by adding independent descriptions of the dominant stimulus category in each region's response profile from participants who are unfamiliar with the theories and research concerning these areas. Moreover, in two separate reliability analyses, we found that these regions exhibited reliable intersubject similarity of response (i.e., ISCs; Hasson et al., 2004; Lerner et al., 2011), as well as good reliability in the selection of scenes and category types in the reverse correlation analysis.

To determine the reliability of the reverse correlation procedure, we used Monte Carlo simulations based on 1000 iterations of partitioning the data in half and comparing the response profiles generated in each half. It was found that the DMPFC demonstrated moderate reliability in the peaks identified by the reverse correlation analysis, though this was generally much less than in visual areas especially at higher statistical thresholds. Despite this, the average distribution of the movie scene categories in the response profile of the DMPFC for both halves of the data was nearly identical. Together, these findings provide evidence for differential category preferences across the four regions studied and lend support to the view that the DMPFC is tuned for the processing of social information.

Over the past several years, the use of naturalistic stimuli in cognitive neuroscience research has steadily increased. A number of studies have relied on natural visual scenes (Wagner et al., 2011b), audio and verbal stories (Yarkoni et al., 2008; Skipper et al., 2009; Speer et al., 2009; Regev et al., 2013), and, perhaps most complex of all, audiovisual narratives in the form of popular motion pictures and television shows (Bartels and Zeki, 2004; Hasson et al., 2004; Moran et al., 2004; Haxby et al., 2011; Wagner et al., 2011a; Byrge et al., 2015; Rapuano et al., 2015). This research has largely confirmed prior results in more constrained experimental settings, demonstrating, for example, that during audiovisual narratives, the lateral and medial fusiform gyrus have been found to respond to scenes of faces (Bartels and Zeki, 2004; Hasson et al., 2004) and places (Hasson et al., 2004). Our findings add to this work by providing evidence that the DMPFC is, in our data, selectively tuned to the presence of social interactions within movie scenes and demonstrated selectivity in its peak re-
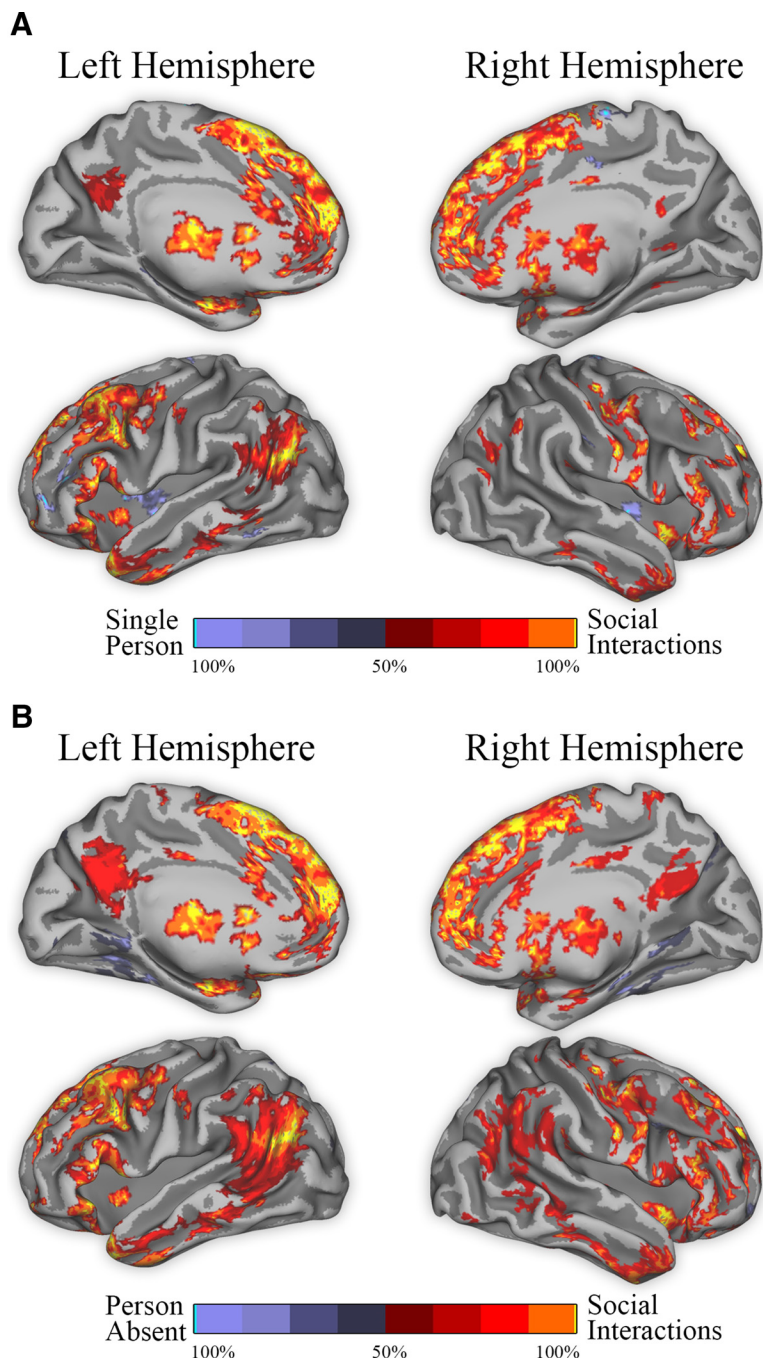
Figure 3. Results of a whole-brain reverse correlation analysis for regions in which a distribution of category preference deviated significantly from the distribution of scene types in the movie ($\chi^2 > 28.94$, $p < 0.05$ corrected). Areas with peaks dominated by social scenes are represented in bright yellow (100% social interactions). Regions with peaks dominated by single person scenes (**A**) and by person absent scenes (**B**) are represented in light blue.

and the passage of time to the traits, goals, and emotions of fictional characters (Gernsbacher et al., 1992; Zwaan and Radvansky, 1998; Rapp et al., 2001). For instance, behavioral work has shown that, when people read passages that do not conform to their expectations of a character's personality (Rapp et al., 2001) or of their current emotional states (Gernsbacher et al., 1992), reading rates slow down, suggesting that readers are maintaining an internal model of a character's personality and current mental states. Consistent with this notion, our results show that DMPFC is primarily responding to movie scenes involving social interactions, which are those scenes in which the audience is most likely to learn about a character's intentions and personalities. In contrast, the troughs contained scenes in which the audience is least likely to find information to update their character models. Prior research has demonstrated a role for the DMPFC in updating impressions of people from behavioral descriptions (Mende-Siedlecki et al., 2013) and we speculate that, during movie viewing, the DMPFC may generate and update character situation models that are built primarily out of observations of those characters' social interactions. In this framework, mental state reasoning is being used to update our inferences about characters' beliefs and intentions, feeding the output of this mentalizing process back into a person knowledge representation for each character and ultimately allowing the audience to generate predictions of a character's behavior. Future work could attempt to capture this process by investigating whether scenes that are more diagnostic of a character's personality differentially recruit the DMPFC during natural viewing.

One limitation of the reverse correlation method used here is that the underlying hemodynamic process is difficult to model, requiring that we make assumptions about the type of hemodynamics contributing to each region's cortical response. In the present study, we assumed a standard time to peak of ~5 s consistent with the canonical hemodynamic response (Miezin et al., 2000). In addition, the present method works best when relying on a priori ROIs. Attempts to perform voxelwise whole-brain mapping of response profiles using this technique would generate an unwieldy amount of data for subsequent analysis and characterization. That said, a comparatively crude characterization of each voxel's response profile was attempted by calculating the response profile to the movie at each voxel and identifying the category membership of peak responses using a simple segmentation of the movie that binned scenes according to whether each frame contained a social interaction, single person, or no persons

sponses and in its troughs, which were conspicuously devoid of social interactions. This selectivity likely reflects the greater amount of information inherent in these scenes that participants can use to perform the types of social cognitive operations long believed to be implemented in this region (for reviews, see Mitchell, 2008; Van Overwalle, 2009; Wagner et al., 2012).

The present findings also have implications regarding theories of narrative comprehension and its functional neuroanatomy. It has been theorized that, to understand narratives, readers spontaneously construct situation models describing the current states of different aspects of a story, from the location of objects

(see Materials and Methods). In this way, an initial characterization of the response profile at each voxel could be determined. Findings from this analysis largely corroborated the ROI analysis demonstrating that social interactions dominate the response profile across the MFPC and temporoparietal junction and temporal poles. In future work, it may be possible to leverage unsupervised learning techniques to define a functional parcellation (Wig et al., 2014; Gordon et al., 2014) based on the similarity of neighboring voxels' response profiles to reduce the problem to a more tractable set of functionally defined brain regions and thereby enable a finer-grained analysis of each functional region's response profile during natural viewing.

Conclusion

As a social species, people spend much of their day engaging in real or mediated social interactions and the better part of their evenings watching social interactions flicker by on television. Therefore, social perception of real or fictional others is, for many, the predominant cognitive mode of everyday life. In the present study, participants engaged in a reasonably naturalistic task (i.e., movie watching) to determine whether the role of the DMPFC in social cognition generalized to more ecologically valid contexts. Using a data-driven reverse correlation technique to visualize the response profile of the DMPFC, we found that this region is particularly tuned to the social features of an audiovisual narrative despite the hundreds of potential stimulus categories present in the movie. Even under natural viewing conditions and in the absence of any explicit task, the DMPFC appeared to track the ebb and flow of social information, thereby confirming the importance of this region for real-world social cognition.

Notes

Supplemental material for this article is available at http://wagnerlab. science/notebooks/revcor2016.html. This material has not been peer reviewed.

References

Bartels A, Zeki S (2004) Functional brain mapping during free viewing of natural scenes. Hum Brain Mapp 21:75–85. CrossRef Medline

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B (Methodological) 57:289–300.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. The Annals of Statistics 29: 1165–1188. CrossRef

Byrge L, Dubois J, Tyszka JM, Adolphs R, Kennedy DP (2015) Idiosyncratic brain activation patterns are associated with poor social comprehension in autism. J Neurosci 35:5837–5850. CrossRef Medline

Connolly AC, Guntupalli JS, Gors J, Hanke M, Halchenko YO, Wu YC, Abdi H, Haxby JV (2012) The representation of biological classes in the human brain. J Neurosci 32:2608–2618. CrossRef Medline

Dal Cin S, Zanna MP, Fong GT (2004) Narrative persuasion and overcoming resistance. In: Resistance and persuasion (Knowles ES, Linn JA, eds). Mahwah, NJ: Erlbaum Associates.

Davis MH (1983) Measuring individual differences in empathy: evidence for a multidimensional approach. Journal of Personality and Social Psychology 44:113–126. CrossRef

Ferrari C, Lega C, Vernice M, Tamietto M, Mende-Siedlecki P, Vecchi T, Todorov A, Cattaneo Z (2016) The dorsomedial prefrontal cortex plays a causal role in integrating social impressions from faces and verbal descriptions. Cereb Cortex 26:156–165. CrossRef Medline

Gernsbacher MA, Goldsmith HH, Robertson RR (1992) Do readers mentally represent characters' emotional states? Cogn Emot 6:89–111. CrossRef Medline

Gobbini MI, Leibenluft E, Santiago N, Haxby JV (2004) Social and emotional attachment in the neural representation of faces. Neuroimage 22: 1628–1635. CrossRef Medline

Gobbini MI, Gentili C, Ricciardi E, Bellucci C, Salvini P, Laschi C, Guazzelli M, Pietrini P (2011) Distinct neural systems involved in agency and animacy detection. J Cogn Neurosci 23:1911–1920. CrossRef Medline

Gordon EM, Laumann TO, Adeyemo B, Huckins JF, Kelley WM, Petersen SE (2016) Generation and evaluation of a cortical area parcellation from resting-state correlations. Cereb Cortex 26:288–303. CrossRef Medline

Hasson U, Honey CJ (2012) Future trends in neuroimaging: neural processes as expressed within real-life contexts. Neuroimage 62:1272–1278. CrossRef Medline

Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R (2004) Intersubject synchronization of cortical activity during natural vision. Science 303: 1634–1640. CrossRef Medline

Haxby JV, Ungerleider LG, Horwitz B, Maisog JM, Rapoport SI, Grady CL (1996) Face encoding and recognition in the human brain. Proc Natl Acad Sci U S A 93:922–927. CrossRef Medline

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P (2001) Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293:2425–2430. CrossRef Medline

Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, Gobbini MI, Hanke M, Ramadge PJ (2011) A common, high-dimensional model of the representational space in human ventral temporal cortex. Neuron 72:404–416. CrossRef Medline

Iacoboni M, Lieberman MD, Knowlton BJ, Molnar-Szakacs I, Moritz M, Throop CJ, Fiske AP (2004) Watching social interactions produces dorsomedial prefrontal and medial parietal BOLD fMRI signal increases compared to a resting baseline. Neuroimage 21:1167–1173. CrossRef Medline

Ishai A, Ungerleider LG, Martin A, Schouten JL, Haxby JV (1999) Distributed representation of objects in the human ventral visual pathway. Proc Natl Acad Sci U S A 96:9379–9384. CrossRef Medline

Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17:4302–4311. Medline

Lerner Y, Honey CJ, Silbert LJ, Hasson U (2011) Topographic mapping of a hierarchy of temporal receptive windows using a narrated story. J Neurosci 31:2906–2915. CrossRef Medline

Mende-Siedlecki P, Cai Y, Todorov A (2013) The neural dynamics of updating person impressions. Soc Cogn Affect Neurosci 8:623–631. CrossRef Medline

Miezin FM, Maccotta L, Ollinger JM, Petersen SE, Buckner RL (2000) Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. Neuroimage 11:735–759. CrossRef Medline

Mitchell JP (2008) Contributions of functional neuroimaging to the study of social cognition. Current Directions in Psychological Science 17: 142–146. CrossRef

Moran JM, Wig GS, Adams RB Jr, Janata P, Kelley WM (2004) Neural correlates of humor detection and appreciation. Neuroimage 21:1055–1060. Medline

Moran JM, Jolly E, Mitchell JP (2012) Social-cognitive deficits in normal aging. J Neurosci 32:5553–5561. CrossRef Medline

Mur M, Ruff DA, Bodurka J, De Weerd P, Bandettini PA, Kriegeskorte N (2012) Categorical, yet graded, single-image activation profiles of human category-selective cortical regions. J Neurosci 32:8649–8662. CrossRef Medline

Rapp DN, Gerrig RJ, Prentice DA (2001) Readers' trait-based models of characters in narrative comprehension. Journal of Memory and Language 45:737–750. CrossRef

Rapuano KM, Huckins JF, Sargent JD, Heatherton TF, Kelley WM (2016) Individual differences in reward and somatosensory-motor brain regions correlate with adiposity in adolescents. Cereb Cortex 26:2602–2611. Medline

Regev M, Honey CJ, Simony E, Hasson U (2013) Selective and invariant neural responses to spoken and written narratives. J Neurosci 33:15978–15988. CrossRef Medline

Skipper JI, Goldin-Meadow S, Nusbaum HC, Small SL (2009) Gestures orchestrate brain networks for language understanding. Curr Biol 19: 661–667. CrossRef Medline

Speer NK, Reynolds JR, Swallow KM, Zacks JM (2009) Reading stories activates neural representations of visual and motor experiences. Psychol Sci 20:989–999. CrossRef Medline

Spiers HJ, Maguire EA (2006) Spontaneous mentalizing during an interac-

tive real world task: An fMRI study. Neuropsychologia 44:1674–1682. CrossRef Medline

Spreng RN, Mar RA, Kim AS (2009) The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. J Cogn Neurosci 21:489–510. CrossRef Medline

Spunt RP, Adolphs R (2015) Folk explanations of behavior: a specialized use of a domain-general mechanism. Psychol Sci 26:724–736. CrossRef Medline

Van Essen DC, Drury HA, Dickson J, Harwell J, Hanlon D, Anderson CH (2001) An integrated software suite for surface-based analyses of cerebral cortex. Journal of the American Medical Informatics Association 8: 443–459. CrossRef Medline

Van Overwalle F (2009) Social cognition and the brain: a meta-analysis. Hum Brain Mapp 30:829–858. CrossRef Medline

Vul E, Lashkari D, Hsieh PJ, Golland P, Kanwisher N (2012) Data-driven functional clustering reveals dominance of face, place, and body selectivity in the ventral visual pathway. J Neurophysiol 108:2306–2322. CrossRef Medline

Wagner DD, Dal Cin S, Sargent JD, Kelley WM, Heatherton TF (2011a) Spontaneous action representation in smokers when watching movie characters smoke. J Neurosci 31:894–898. CrossRef Medline

Wagner DD, Kelley WM, Heatherton TF (2011b) Individual differences in the spontaneous recruitment of brain regions supporting mental state

understanding when viewing natural social scenes. Cereb Cortex 21: 2788–2796. CrossRef Medline

Wagner DD, Haxby JV, Heatherton TF (2012) The representation of self and person knowledge in the medial prefrontal cortex. Wiley Interdiscip Rev Cogn Sci 3:451–470.

Wheatley T, Milleville SC, Martin A (2007) Understanding animate agents distinct roles for the social network and mirror system. Psychol Sci 18: 469–474. CrossRef Medline

Wig GS, Laumann TO, Petersen SE (2014) An approach for parcellating human cortical areas using resting-state correlations. Neuroimage 93: 276–291. CrossRef Medline

Wu MC, David SV, Gallant JL (2006) Complete functional characterization of sensory neurons by system identification. Annu Rev Neurosci 29: 477–505. CrossRef Medline

Yarkoni T, Speer NK, Zacks JM (2008) Neural substrates of narrative comprehension and memory. Neuroimage 41:1408–1425. CrossRef Medline

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD (2011) Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8:665–670. CrossRef Medline

Zaki J, Ochsner K (2009) The need for a cognitive neuroscience of naturalistic social cognition. Ann N Y Acad Sci 1167:16–30. CrossRef Medline

Zaki J, Weber J, Bolger N, Ochsner K (2009) The neural bases of empathic accuracy. Proc Natl Acad Sci U S A 106:11382–11387. CrossRef Medline

Zwaan RA, Radvansky GA (1998) Situation models in language comprehension and memory. Psychol Bull 123:162–185. CrossRef Medline